

## PATENT ABSTRACTS OF JAPAN

(11) Publication number : 2004-151757

(43) Date of publication of application : 27.05.2004

(51)Int.Cl. G06F 17/21

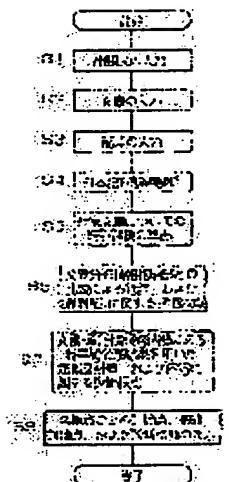
(21)Application number : 2002-313004 (71)Applicant : RICOH CO LTD  
(22)Date of filing : 28.10.2002 (72)Inventor : ISHIOKA TSUNENORI  
KAMEDA MASAYUKI

(54) SENTENCE EVALUATING AND SCORING DEVICE, PROGRAM, AND STORAGE MEDIUM

(57) Abstract:

**PROBLEM TO BE SOLVED:** To enable a user to freely set and change weighting (allocation of points) for each view point.

**SOLUTION:** An object sentence is evaluated and scored from a plurality points of view respectively and each of the evaluated score is integrated (step S4 to S7). The result is displayed on a display device (step S8). Setting of weighting for each of the plurality of points of view can be changed by accepting operation from the user (step S3).



(19) 日本国特許庁(JP)

## (12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-151757

(P2004-151757A)

(43) 公開日 平成16年5月27日(2004.5.27)

(51) Int.Cl.<sup>7</sup>  
G06F 17/21F 1  
G06F 17/21 550Aテーマコード(参考)  
5B009

審査請求 未請求 請求項の数 11 O L (全 14 頁)

(21) 出願番号	特願2002-313004 (P2002-313004)	(71) 出願人	000006747 株式会社リコー 東京都大田区中馬込1丁目3番6号
(22) 出願日	平成14年10月28日 (2002.10.28)	(74) 代理人	100101177 弁理士 柏木 慎史
		(74) 代理人	100102130 弁理士 小山 尚人
		(74) 代理人	100072110 弁理士 柏木 明
		(72) 発明者	石岡 恒憲 東京都目黒区駒場2-19-23 独立行政法人 大学入試センター内
		(72) 発明者	亀田 雅之 東京都大田区中馬込1丁目3番6号 株式会社リコー内
			F.ターム(参考) 5B009 QA02 QA04 QA14

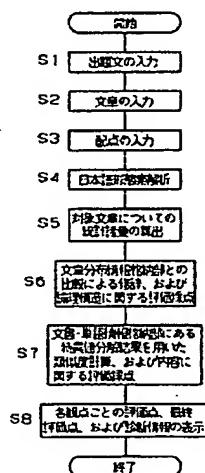
(54) 【発明の名称】文章評価採点装置、プログラム及び記憶媒体

## (57) 【要約】

【課題】ユーザは各観点にかかる重み付け(配点)を自由に設定し、変更することができるようとする。

【解決手段】対象となる文章を複数の観点からそれぞれ評価採点し、その各評価採点を統合する(ステップS4～S7)。その結果は、表示装置に表示する(ステップS8)。複数の各観点の重み付けの設定は、ユーザから受付けることで変更可能である(ステップS3)。

【選択図】 図4



**【特許請求の範囲】****【請求項 1】**

対象となる文章を複数の観点からそれぞれ評価採点し、その各評価採点を統合する文章評価採点手段と、

この評価採点の結果を出力する出力手段と、

前記各観点の重み付けの設定をユーザから受付ける重み付け設定手段と、  
を備えている文章評価採点装置。

**【請求項 2】**

対象となる文章を 1 又は複数の観点からそれぞれ評価採点し、その各評価採点を統合する文章評価採点手段と、

10

この評価採点の結果を出力する出力手段と、

を備え、

前記文章評価採点手段は、前記複数の観点のうち 1 又は複数のものをそれぞれ示す複数の指標であって前記対象となる文章のものと、前記各指標であって複数の文書から収集したもののそれぞれの分布の情報とをそれぞれ比較することで前記評価採点を行なう、  
文章評価採点装置。

**【請求項 3】**

前記文章評価採点手段は、前記対象となる文章の前記各指標が前記各指標の分布に対してそれぞれ統計学的に外れているか否かを判断して前記評価採点を行なう、請求項 2 に記載の文章評価採点装置。

20

**【請求項 4】**

前記文章評価採点手段は、前記対象となる文章の前記各指標の前記各指標の分布に対する偏差値によって前記評価採点を行なう、請求項 2 に記載の文章評価採点装置。

**【請求項 5】**

前記文章評価採点手段は、前記文章についての修辞、論理構成及び内容のうち少なくとも一つを前記観点としている、請求項 1 ~ 4 の何れかの一に記載の文章評価採点装置。

30

**【請求項 6】**

前記文章評価採点手段は、前記修辞の観点を示す前記指標として、読みやすさ、語彙の多様性、ビッグワードの割合及び受動態の文の割合、の各指標のうちの少なくとも一つ以上を用いる、請求項 5 に記載の文章評価採点装置。

**【請求項 7】**

前記文章評価採点手段は、前記読みやすさの指標として、文字数、文数及び文の長さ、文内文数及び文内文の長さ、文内文中の句数、文字種の比率、連体修飾文（埋め込み文）の数、並びに、連用修飾文の並びの数、の各細目指標のうちの少なくとも一つ以上を用いる、請求項 5 に記載の文章評価採点装置。

40

**【請求項 8】**

前記文章評価採点手段は、前記論理構成の観点を示す前記指標として、文章の接続関係を示す順接及び逆接の接続詞の数、並びに、文章の接続関係を示す順接及び逆接の接続詞の出現パターンのうちの少なくとも一つ以上を用いる、請求項 5 に記載の文章評価採点装置

。

**【請求項 9】**

前記文章評価採点手段は、前記内容の観点の前記評価採点を、複数の文書から作成した文書一単語行列を特異値分解した結果を利用して、前記対象となる文書の作成意図となる出題文と前記対象となる文章の類似度を計算し、この結果を利用して行なう、請求項 5 に記載の文章評価採点装置。

**【請求項 10】**

請求項 1 ~ 9 の何れかの一に記載の文章評価採点装置の前記各手段の機能をコンピュータに実行させる、コンピュータに読み取り可能なプログラム。

50

**【請求項 11】**

請求項 10 に記載のプログラムを記憶している記憶媒体。

## 【発明の詳細な説明】

## 【0001】

## 【発明の属する技術分野】

本発明は、文章評価採点装置、プログラム及び記憶媒体に関する。

## 【0002】

## 【従来の技術】

文章の評価方法に関する発明として、特許文献1、2には、文章の修辞に関する評価に関する技術について開示されているが、文章の論理構成や内容の適切さについては評価されていない。特許文献3には、文書の種類に応じた評価を行なう文書評価方式について開示されており、また、特許文献4には、文章要約における重要文を抽出するためにニューラルネットによる学習を利用する技術について開示されている。10

## 【0003】

## 【特許文献1】

特許第2609650号公報

## 【特許文献2】

特許第2615121号公報

## 【特許文献3】

特開平6-12410号公報

## 【特許文献4】

特開平5-174019号公報

20

## 【非特許文献1】

Cooper, P. L. : The assessment of writing ability: a review of research, GRE Board Research Report, GREB No. 82-15R. 1984

## 【0004】

## 【発明が解決しようとする課題】

アメリカにおいては、近年、コンピュータによる小論文の自動採点の研究が精力的に行なわれており、その中でもアメリカのテスト機関であるETS (Educational Testing Service) が開発し、現在はその補助機関である ETS Technologies に拡張開発、および運用が移管されている e-rater は最も有名で、経営大学院 (ビジネススクール) の入学試験であるGMAT (Graduate Management Admission Test) における小論文の採点に用いられている。30

## 【0005】

小論文試験においては、実施者は受験者のある種の能力が答案に反映していることを期待しているわけだが、その得点結果には、様々な要因が複雑に関与している。前述の非特許文献1によると、「小論文がWriting Abilityを測定しているものと考えると、その得点に関して誤差要因として働くものには、書き手 (writer)、題目 (topic)、形式 (mode)、制限時間 (time-limit)、テスト状況 (examination situation)、そして評定者 (rater) がある」という。これらの大部分はいわゆる「試験」に共通している要因であるが、特に「評定者」の要因は小論文においては決定的なものである。40

## 【0006】

他にも小論文試験では、得点に影響を与える以下のような多くの要因が存在し、それらについての多くの研究がある。

- ・文字の巧拙 (文字の上手さ、綴りの正確性)
- ・評定の系列的効果 (ある小論文の評定が答案の中で何番目に行なわれたか)
- ・課題選択 (異なる課題に基づいて書かれた小論文をどう評価するか)
- ・その他種々の誤差要因 (書き手の性別、人種など)

e-rater は、このような誤差要因を排除し、試験採点における公平性を担保する。50

そのため、e-raterは以下の3つの観点から小論文を評定する。

構造：文法の多様性、すなわちフレーズや文節、および文の配列が多様な構造で表現されているか

組織化：アイディアが理路整然と表現されているか。たとえば修辞的な表現、あるいは文や節の間の論理的な接続法が用いられているか

内容：トピックに関連した語彙が用いられているか

e-raterでは、この3つの観点を示すと考えられる指標を選び、人間による膨大な数の採点結果とできるだけ一致するように線形回帰を行ない、それぞれの指標にかかる回帰係数を求めている。

#### 【0007】

10

これら3つの観点（構造、組織化、内容）が、論文評価をおこなうための重要な視点であることは疑いがないであろうが、これら3つの観点の比率をどのようにするかは、いわゆる採点基準をどのように定めるかに他ならず、評価すべき論文の性質や採点者のビジョンに大きく依存し、通常は一定ではない。もちろん、先行研究の成果を踏まえ、通常は妥当と考えられる配点を既定値とはするが、この配点は必要に応じてユーザが自由に変更できることが望ましい。

#### 【0008】

20

また、e-raterは、人間の採点にできるだけ近づけるというアプローチをとるために、採点対象となる小論文においてどの指標が優れているのか、あるいは劣っているのかの判断を下すことが困難である。

#### 【0009】

30

本発明の目的は、ユーザは各観点にかかる重み付け（配点）を自由に設定し、変更することができるようになることである。

#### 【0010】

本発明の別の目的は、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的かつ自動的に評価採点を行なうことである。

#### 【0011】

##### 【課題を解決するための手段】

請求項1に記載の発明は、対象となる文章を複数の観点からそれぞれ評価採点し、その各評価採点を統合する文章評価採点手段と、この評価採点の結果を出力する出力手段と、前記各観点の重み付けの設定をユーザから受付ける重み付け設定手段と、を備えている文章評価採点装置である。

#### 【0012】

したがって、ユーザは各観点にかかる重み付け（配点）を自由に設定し、変更することができる。

#### 【0013】

請求項2に記載の発明は、対象となる文章を1又は複数の観点からそれぞれ評価採点し、その各評価採点を統合する文章評価採点手段と、この評価採点の結果を出力する出力手段と、を備え、前記文章評価採点手段は、前記複数の観点のうち1又は複数のものをそれぞれ示す複数の指標であって前記対象となる文章のものと、前記各指標であって複数の文書から収集したもののそれぞれの分布の情報をそれぞれ比較することで前記評価採点を行なう。

40

#### 【0014】

したがって、観点ごとに設けられた各指標と、予め収集した多数の優れた文書における指標の分布とを比較することで評価採点を行なうので、指標ごとの優劣の判定が容易であり、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的かつ自動的に評価採点を行なうことができる。また、この判定結果をコメントとして出力することができる。

#### 【0015】

請求項3に記載の発明は、請求項2に記載の文章評価採点装置において、前記文章評価採

50

点手段は、前記対象となる文章の前記各指標が前記各指標の分布に対してそれぞれ統計学的に外れているか否かを判断して前記評価採点を行なう。

【0016】

したがって、対象となる文章の指標が、予め収集した多数の優れた文書における指標の分布に対して統計学的に外れているか否かで評価採点を行なうので、指標ごとの優劣の判定が容易であり、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。また、この判定結果をコメントとして出力することができる。

【0017】

請求項4に記載の発明は、請求項3に記載の文章評価採点装置において、前記文章評価採点手段は、前記対象となる文章の前記各指標の前記各指標の分布に対する偏差値によって前記評価採点を行なう。10

【0018】

したがって、対象となる文章の指標の、予め収集した多数の優れた文書における指標の分布に対する偏差値によって評価採点を行なうので、指標ごとの優劣の判定が容易であり、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。

【0019】

請求項5に記載の発明は、請求項1～4の何れかの一に記載の文章評価採点装置において、前記文章評価採点手段は、前記文章についての修辞、論理構成及び内容のうち少なくとも一つを前記観点としている。20

【0020】

したがって、文章についての修辞、論理構成及び内容のうち少なくとも一つを観点として、観点ごとに設けられた各指標と、予め収集した多数の優れた文書における指標の分布とを比較することで評価採点を行ない、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。

【0021】

請求項6に記載の発明は、請求項5に記載の文章評価採点装置において、前記文章評価採点手段は、前記修辞の観点を示す前記指標として、読みやすさ、語彙の多様性、ビッグワードの割合及び受動態の文の割合、の各指標のうちの少なくとも一つ以上を用いる。30

【0022】

したがって、修辞の観点を示す指標として、読みやすさ、語彙の多様性、ビッグワードの割合及び受動態の文の割合、の各指標のうちの少なくとも一つ以上を用いて、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。

【0023】

請求項7に記載の発明は、請求項5に記載の文章評価採点装置において、前記文章評価採点手段は、前記読みやすさの指標として、文字数、文数及び文の長さ、文内文数及び文内文の長さ、文内文中の句数、文字種の比率、連体修飾文（埋め込み文）の数、並びに、連用修飾文の並びの数、の各細目指標のうちの少なくとも一つ以上を用いる。40

【0024】

したがって、読みやすさの指標として、文字数、文数及び文の長さ、文内文数及び文内文の長さ、文内文中の句数、文字種の比率、連体修飾文（埋め込み文）の数、並びに、連用修飾文の並びの数、の各細目指標のうちの少なくとも一つ以上を用いることで、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。

【0025】

請求項8に記載の発明は、請求項5に記載の文章評価採点装置において、前記文章評価採点手段は、前記論理構成の観点を示す前記指標として、文章の接続関係を示す順接及び逆接の接続詞の数、並びに、文章の接続関係を示す順接及び逆接の接続詞の出現パターンの50

うちの少なくとも一つ以上を用いる。

【0026】

したがって、論理構成の観点を示す指標として、文章の接続関係を示す順接及び逆接の接続詞の数、並びに、文章の接続関係を示す順接及び逆接の接続詞の出現パターンのうちの少なくとも一つ以上を用いることで、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的につつ自動的に評価採点を行なうことができる。

【0027】

請求項9に記載の発明は、請求項5に記載の文章評価採点装置において、前記文章評価採点手段は、前記内容の観点の前記評価採点を、複数の文書から作成した文書一単語行列を特異値分解した結果を利用して、前記対象となる文書の作成意図となる出題文と前記対象となる文章の類似度を計算し、この結果を利用して行なう。

10

【0028】

したがって、文章の内容について、与えられた出題文との意味的な近さを測定することで評価採点を行なうので、小論文試験において採点に影響を与える様々な誤差要因を排除して、客観的につつ自動的に評価採点を行なうことができる。

【0029】

請求項10に記載の発明は、請求項1～9の何れかの一に記載の文章評価採点装置の前記各手段の機能をコンピュータに実行させる、コンピュータに読み取り可能なプログラムである。

【0030】

したがって、本プログラムをコンピュータに実行させることにより、請求項1～9の何れかの一に記載の発明と同様の作用、効果を奏する。

20

【0031】

請求項11に記載の発明は、請求項10に記載のプログラムを記憶している記憶媒体である。

【0032】

したがって、記憶しているプログラムにより請求項10に記載の発明と同様の作用、効果を奏する。

【0033】

【発明の実施の形態】

30

本発明の一実施の形態について説明する。

【0034】

図1は、本実施の形態である文章評価採点装置1の電気的な接続を示すブロック図である。図1に示すように、文章評価採点装置1は、PCなどのコンピュータであり、各種演算を行ない、文章評価採点装置1の各部を集中的に制御するCPU2と、各種のROMやRAMからなるメモリ3とが、バス4で接続されている。

【0035】

バス4には、所定のインターフェイスを介して、ハードディスクなどの磁気記憶装置5と、マウスやキーボードなどで構成される入力装置6と、LCDやCRTなどの表示装置7と、光ディスクなどの記憶媒体8を読み取る記憶媒体読み取り装置9とが接続され、また、インターネットなどのネットワーク10と通信を行なう所定の通信インターフェイス11が接続されている。なお、記憶媒体8としては、CDやDVDなどの光ディスク、光磁気ディスク、フレキシブルディスクなどの各種方式のメディアを用いることができる。また、記憶媒体読み取り装置9は、具体的には記憶媒体8の種類に応じて光ディスクドライブ、光磁気ディスクドライブ、フレキシブルディスクドライブなどが用いられる。

40

【0036】

磁気記憶装置5には、この発明のプログラムを実現する文章評価採点プログラムが記憶されている。この文章評価採点プログラムは、記憶媒体8から記憶媒体読み取り装置9により読み取るか、あるいは、インターネットなどのネットワーク10からダウンロードするなどして、磁気記憶装置5にインストールしたものである。このインストールにより文章評価採

50

点装置1は動作可能な状態となる。なお、この文章評価採点プログラムは、所定のOS上で動作するものであってもよい。また、特定のアプリケーションソフトの一部をなすものであってもよい。

【0037】

図2に示すように、この文章評価採点装置1をサーバコンピュータ14として実施し、このサーバコンピュータ14と端末装置12とをネットワーク13を介して接続して、端末装置12からサーバコンピュータ14を操作できるようにしてもよい。この場合に、端末装置12は、パーソナルコンピュータ、携帯情報端末(PDA)、携帯電話などの情報処理装置として実施することができる。また、ネットワーク13は、無線、有線及び放送波のいずれを用いたものでもよく、例えば、LAN、WAN、インターネット、アナログ電話網、デジタル電話網(ISDN)、PHS(パーソナルハンディホンシステム)網、携帯電話網、衛星通信網などを利用することができる。10

【0038】

以下では、文章評価採点プログラムに基づいて、文章評価採点装置1が行なう処理の内容について説明する。

【0039】

図3は、文章評価採点装置1の機能ブロック図である。出題文入力部21は、評価採点の対象となる文章の作成意図となる出題文の入力を受付け、これを本文章評価採点装置(のメモリ3、磁気記憶装置5等の記憶媒体)に記憶する。文章入力部22は、評価採点の対象となる対象となる文章の入力を受付け、これを本文章評価採点装置1(のメモリ3、磁気記憶装置5等の記憶媒体)に記憶する。配点入力部23は、評価採点における各観点の配点(重み付け)の設定、修正の入力を受けて、本文章評価採点装置1(のメモリ3、磁気記憶装置5等の記憶媒体)に格納する。20

【0040】

文章評価採点部24は、出題文入力部21、文章入力部22及び配点入力部23で入力したデータに基づき、各観点の採点結果を算出し、これらを統合することで文章評価点を与える。

【0041】

なお、この文章評価採点部24は、詳細には、日本語形態素解析を行なう形態素解析部25を含み、文章入力部22で入力された文章に対して、文法規則が予め格納されてある文法規則格納部26と、辞書を格納している辞書格納部27とを用いて、文章評価を行なうための各指標についての統計諸量を計算しておく。30

【0042】

文章評価採点部24では、文章分布情報格納部28に予め収集してある多数の優れた文章についての修辞及び論理構造に関する各種指標の分布情報を格納しておき、入力された文章の統計量と文章分布情報格納部28に格納された同指標の分布とを比較することで、評価採点を行なう。比較の方法としては、入力された文章の統計量が文章分布情報格納部28に格納された同指標の分布において、統計学的に外れ値とみなすことができるか否かによって判断をする。

【0043】

また、文章の評価採点においては、対象文章が与えられた出題文に対して相応しいものか評価することが不可欠である。このために文書-単語情報格納部29に予め多数の文書から作成した文書-単語行列を特異値分解した結果を格納しておき、この結果を用いて、出題文入力部21で入力した出題文と対象文章の類似度を計算し、これをを利用して評価採点を行なう。

【0044】

文章評価採点表示部30では、文章評価採点部24の結果をもとに、それぞれの観点ごとの評価点と、そこから得られる最終評価点を表示装置7に表示するとともに、対象文章に対する診断情報を表示する。

【0045】

10

20

30

40

50

以下では、文章評価採点装置1の各部で行なう処理の詳細な内容について説明する。

【0046】

図4は、かかる処理を説明するフローチャートである。図4に示すように、出題文入力部21で出題文の入力を受付け(ステップS1)、文章入力部22で出題文の入力を受付け(ステップS2)。コンピュータで処理可能な日本語コードには、大別してパーソナルコンピュータでよく用いられるシフトJIS、UNIXマシンでよく用いられるEUC(Extended Unix Code)、通信でよく用いられる7ビットのJISの3つが知られているが(UNIX、UNIXは登録商標である)、ステップS1、S2における入力は、このうちのいずれで入力してもよい。入力された文章は、文章評価採点装置1において全てEUCに変換して処理する。

10

【0047】

また、配点入力部23で、配点の入力を受付け、設定手段を実現する(ステップS3)。これは、評価対象となる文章は3つの観点、すなわち、修辞、論理構成、内容の少なくとも1つ以上の観点で評価されるが、それぞれの観点に対する配点(重み付け)をここで指定する。これは、合計10点満点で、あらかじめ適切な配点として、デフォルトとして、例えば、修辞、論理構成、内容をそれぞれ5、2、3の配点とするが、ユーザはこの配点を自由に変更することができる。なお、ここで修辞、論理構成、内容をそれぞれ5、2、3とする配点例は、“渡部洋、平由実子、井上俊哉：小論文評価データの解析、東京大学教育学部紀要、第28巻、143-164. 1988”の研究結果に基づいている。

20

【0048】

そして、入力を受付けた出題文と評価対象となる文章に対して、形態素解析部25により日本語形態素解析を行なう(ステップS4)。ステップS4～S7は、文章評価採点手段を実現するものである。これは、前述のように文法規則格納部26、辞書格納部27を用いて行なう。日本語形態素解析については周知の技術であり、さまざまな形態素解析系を利用できるが、たとえば、奈良先端科学技術大学院大学 松本研究室の茶筌(ちゃせん、“http://chasein.naist-nara.ac.jp”を参照)と、それが用いる日本語文法規則と日本語辞書を使用することができる。

30

【0049】

次に、この形態素結果を用いて、文章評価採点部24が対象文書に関する各指標についての統計諸量を算出する(ステップS5)。

【0050】

(1) 日本語の「修辞」を示す統計量には、“前川守：文章を科学する、1000万人のコンピュータ科学3、岩波書店、1995、1章pp. 1-56”、“長尾真(編)：自然言語処理、岩波講座ソフトウェア科学15、岩波書店、1996、2. 1節pp. 13-35参照”によれば以下が挙げられる。

40

【0051】

▲1▼. 文章の読みやすさを示すメトリクスとして

- ・文の長さの中央値、最大値
- ・句の長さの中央値、最大値
- ・句中における文節数の中央値、最大値
- ・漢字／カナの割合
- ・連体修飾(埋め込み文)の数
- ・連用形や接続助詞の句の並びの最大値

▲2▼. 語彙の多様性を示すメトリクスとして

- ・ユール(Yule)のK特性値

▲3▼. 文章の専門性を示すメトリクスとして

- ・ビッグワード(長くて難しい語)の発生割合

▲4▼. いわゆる良い文章を示すメトリクスとして

- ・受動態の文の割合

50

【0052】

(2) また、「論理構造」を把握するために、「順接」および「逆接」を示す接続詞の数と、その出現パターンを記録する。ここで「順接」という語はやや広い意味で用いており、議論の流れが変わらない接続構造一般を指している。

【0053】

日本語の接続構造としては、“野矢茂樹：論理トレーニング、産業図書、第1章および第2章、p.p. 3-26 参照”によれば、以下が挙げられる。

【0054】

▲1▼. 「順接」の接続構造

・付加：主張を加える接続関係である。典型的には「そして」、「しかも」、「むしろ」など。  
10

・解説：典型的には「すなわち」、「つまり」、「言い換えれば」、「要約すれば」といった接続表現で表される接続関係である。

・論証：理由と帰結の関係を示す。理由を示す典型的な接続表現には、「なぜなら」、「その理由は」などがあり、帰結を示すものとしては、「それゆえ」、「したがって」、「だから」、「つまり」などがある。

・例示：典型的には「たとえば」で表される接続関係であり、具体例による解説、ないし論証としての構造をもつ。

【0055】

▲2▼. また、「逆接」の接続構造には以下のものがある。

・転換：ある主張Aに対して対立する主張Bが続けられるとき、Bの方に言いたいことが 20 位置する接続関係をいう。一般に、「AだがB」、「A、しかしB」という表現をとる。

・譲歩：転換の一種とみることもできるが、譲歩の場合は対話的構造が現われる。典型的には、「たしかに」、「もちろん」などである。

・対比：典型的には、「一方」、「他方」、「それに対して」といった接続表現で表される接続関係である。

【0056】

以上のように、各指標についての統計諸量を算出した後（ステップS5）、得られた各指標についての統計諸量を、文章分布情報格納部28に格納された同指標の分布とを比較することで評価採点を行なう（ステップS6）。

【0057】

文章分布情報格納部28には、「修辞」を示す前述のメトリクスに対して、優れた文章のそれらの統計量の分布として、たとえば、新聞記事データベースに納められている社説、あるいはコラムについて得て、格納する。

【0058】

これらメトリクスの分布のほとんどは、左右非対象の歪んだ分布となるが、この分布を理想とする小論文についての分布とみなす。得られた統計量が、この理想とする分布において外れ値となった場合に、そのメトリクスにおいて「適当でない」と判断し、割り当てられた配点を減じ、またその旨をコメントとして出力する。

40

【0059】

外れ値は、四分範囲の1.5倍を越えるデータとする。採点において、ブレークダウンした各メトリクスの比重は同等とする。唯一の例外は「語彙の多様性」の尺度であり、これだけがその重み付けを2倍にする。これは、この項目が修辞だけでなく、内容にも関与する指標であると考えられることによる。

【0060】

例えば、評価に用いられている文章中にある各文の長さの中央値が84文字を超えていたらば、文が総じて（平均的に）長いと判断し、このメトリクスに割り当てられた配点0.2点を減じ、「文が総じて（平均的に）長いです」という出力を出す。文の長さの中央値が50文字を超えたときは、0.1点を減じ、「文が総じて（平均的に）少し長いです

50

」という出力を出す。

【0061】

また、語彙の多様性を示すユールのK特性値の値が172よりも大きければ、このメトリクスに割り当てられた配点0.4点を減じ、「語彙の多様性が不足しています」という出力を出す。116よりも大きくて172以下のときは、0.2点を減じ、「語彙の多様性がやや不足しています」という出力を出す。

【0062】

「論理構造」についても、ここでは新聞の社説に現われる接続関係を示す句を全て抜き出し、これを前述の順接、逆接、各4通り、計8通りに排他的に分類する。

【0063】

10

本例では、採点する小論文の談話 (discourse, 議論のかたまり) に対して接続関係を示すラベルを付加し、これらの個数をカウントすることで議論がよく掘り下げられているか否かを判断する。

【0064】

個数についても、修辞同様、新聞の社説で学習し、模範とする分布において外れ値となつた場合に配点を減ずることとする。例えば、800字の論文で順接や逆説を示す接続詞が全くなければ1点を減じ、「議論の接続が不十分であるように見受けられます」という出力を行なう。接続詞が1個であるならば、0.5点を減じ、「議論の接続がやや不十分であるように見受けられます」という出力を行なう。

【0065】

20

また、これら接続関係の出現パターンが、社説のそれに比べて特異でないかを判断する。そのために、順接と逆接の出現パターンについて、トライグラムモデル（“北研二：確率的言語モデル、言語と計算4、東京大学出版会、1999、3章、pp. 57-99”を参照）を考える。

【0066】

一般にNグラムモデルは、確率有限オートマトンによって表現することができる。オートマトンの各状態は、トライグラムモデルにおいては、長さ2の記号列によりラベル付けされる。記号の集合は、“ $\Sigma = \{a : \text{順接}, b : \text{逆接}\}$ ”である。

【0067】

30

図5は、統計量として得た、各状態遷移の条件付き出力確率の例である。“-”は何もないことを示す。初期状態は“--”である。例えば、“ $P(a | --)$ ”は初期状態で最初に“a : 順接”が出現する確率をいう。

【0068】

これより、論文中の $\{a : \text{順接}\}$ と $\{b : \text{逆接}\}$ の出現パターンに対する生起確率が、第3図に示す条件付き確率の積をとることで得ることができる。例えば、 $\{a, b, a, a\}$ の出現パターンに対する生起確率 $p$ は、“ $0.44 \times 0.52 \times 0.55 \times 0.28 = 0.035$ ”となる。

【0069】

40

一方、事前情報なしに、 $\{a : \text{順接}\}$ の出現する確率は0.47で、 $\{b : \text{逆接}\}$ の出現する確率は0.53であるから、順接が3回と逆接が1回出現したときの、事前情報が与えられていないという条件のもとでの与えられた出現パターンの生起確率 $q$ は、“ $0.47^3 \times 0.53 = 0.055$ ”となる。

【0070】

この例のように、事前情報のない方がその生起確率が大きくなるとき、順接と逆接の出現パターンは特異であると考え、議論の接続に割り当てられた配点、たとえば1点を減ずる。

【0071】

50

以上のようにしてステップS6の処理を行った後、ステップS1で入力された出題文と、ステップS2で入力された文章との内容的な近さ、すなわち類似度を計算することで、書かれている文章が出題文に対して適切な内容になっているかを、文書-単語情報格納部2

9に格納された特異値分解結果を利用し、文章評価採点部24が「内容」についての評価採点を行なう（ステップS7）。

【0072】

この類似度を計算するために、文書検索の国際的評価会TREC (Text Retrieval Conference) などでその有用性が主張されているLSI (Late n t Semantic Indexing) を用いる。このLSIは、予め十分に多くの文書に出現する単語の頻度を表した、“ $t \times d$ の行列X”（但し、 $t$ は単語数、 $d$ は文書数）を特異値分解することから始まる。

【0073】

得られた特異値ベクトルの特異値の大きい方から $k$ 番目までとり、これを対角要素とする<sup>10</sup>対角行列をSとする。それに応じて、 $k$ 列までを抜き出した左右の特異値分解行列をそれぞれT, Dとすれば、

【数1】

$$\hat{X} = T S D' \quad \dots \dots (1)$$

となり、

【数2】

$$\hat{X} \quad \dots \dots (2)$$

は、Xの近似となる。

20

ここで、Tは $t \times k$ 行列、Sは $k \times k$ の正方対角行列、D'は $k \times d$ 行列である（'は転置を示す）。言語データの場合、経験的に $k$ は50～100程度にすればよい。

【0074】

採点される小論文eは、形態素解析によりその小論文が含む<sup>11</sup>次元の単語ベクトルx<sub>e</sub>で表現することができ、これを用いて、文書空間Dの行に対応する、“ $1 \times k$ ”の文書ベクトル

【数3】

$$d = x_e T S^{-1} \quad \dots \dots (3)$$

を導くことができる。出題文qについても同様に $k$ 次元ベクトルd<sub>q</sub>を得ることができる<sup>30</sup>。これより、両文書の近似度r<sub>(d<sub>e</sub>, d<sub>q</sub>)</sub>を、(3)式のように両文書ベクトルがなす角の余弦で与えることができる。

【0075】

【数4】

$$r(d_e, d_q) = (d_e, d_q) / \|d_e\| / \|d_q\| \quad \dots \dots (4)$$

【0076】

(3)式はd<sub>e</sub>とd<sub>q</sub>が正規分布に従うという仮定のもとで、統計学でいうところの相関係数に一致する。したがって本例では、ここで与えられるrを「内容」に割り当てられた配点を乗ずることで、「内容」に対する評点とすることとした。例えば、内容に対する配点が3で“近似度r=0.8”ならば、「内容」に対する評点は2.4である。rは理論的には負の値を取りうるが、その下限を0にすることは妥当であると考えられる。

40

【0077】

以上のようにして、文章評価採点部24が「内容」についての評価採点を行なった後（ステップS7）、得られた評価についての各観点（「修辞」、「論理構成」、「内容」）についての評価採点結果を統合することで最終得点を計算し、これらの採点結果と各観点での評価から得られる文章についての診断情報を文章評価採点表示部30が表示し、出力手段を実現する（ステップS8）。

【0078】

例えば、「修辞」の評点が5点満点中3.0点、「論理構成」の評点が2点満点中1.0<sup>50</sup>

点、「内容」の評点が3点満点中2.9点であるならば、総合の評点は10点満点中6.9点となる。また、各観点についての表示されうる診断情報は、図6に示す通りである。同じ評点でも異なった診断情報を表示することができる。

【0079】

以上のような処理によれば、文章入力部22で入力された文章の修辞及び論理構成について、文章分布情報格納部28に格納された予め収集した多数の優れた文書における同指標の分布に対して、文章評価採点部24で統計学的に外れているか否かで評価採点を行ない、また文章の内容については文書－単語情報格納部29に格納された特異値分解結果を利用することで、与えられた出題文との意味的な近さを測定することで評価採点を行ない、文章評価採点表示部30で最終的な評点を表示することができる。これにより、客観的にかつ自動的に文書の評価採点を行なうことが可能になる。<sup>10</sup>

【0080】

また、配点入力部23では、必要に応じてユーザが採点基準となる3つの観点に対する配点を変更することができる。

【0081】

さらに、文章評価採点表示部30では、最終的な点だけでなく、図6に例示するような診断情報も提示することができる。

【0082】

【発明の効果】

請求項1に記載の発明は、ユーザは各観点にかかる重み付け（配点）を自由に設定し、変更することができる。<sup>20</sup>

【0083】

請求項2に記載の発明は、観点ごとに設けられた各指標と、予め収集した多数の優れた文書における指標の分布とを比較することで評価採点を行なうので、指標ごとの優劣の判定が容易であり、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。また、この判定結果をコメントとして出力することができる。

【0084】

請求項3に記載の発明は、請求項2に記載の発明において、対象となる文章の指標が、予め収集した多数の優れた文書における指標の分布に対して統計学的に外れているか否かで評価採点を行なうので、指標ごとの優劣の判定が容易であり、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。また、この判定結果をコメントとして出力することができる。<sup>30</sup>

【0085】

請求項4に記載の発明は、請求項3に記載の発明において、対象となる文章の指標の、予め収集した多数の優れた文書における指標の分布に対する偏差値によって評価採点を行なうので、指標ごとの優劣の判定が容易であり、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。

【0086】

請求項5に記載の発明は、請求項1～4の何れかの一に記載の発明において、文章についての修辞、論理構成及び内容のうち少なくとも一つを観点として、観点ごとに設けられた各指標と、予め収集した多数の優れた文書における指標の分布とを比較することで評価採点を行ない、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。<sup>40</sup>

【0087】

請求項6に記載の発明は、請求項5に記載の発明において、修辞の観点を示す指標として、読みやすさ、語彙の多様性、ビッグワードの割合及び受動態の文の割合、の各指標のうちの少なくとも一つ以上を用いて、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。<sup>50</sup>

**【0088】**

請求項7に記載の発明は、請求項5に記載の発明において、読みやすさの指標として、文字数、文数及び文の長さ、文内文数及び文内文の長さ、文内文中の句数、文字種の比率、連体修飾文（埋め込み文）の数、並びに、連用修飾文の並びの数、の各細目指標のうちの少なくとも一つ以上を用いることで、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。

**【0089】**

請求項8に記載の発明は、請求項5に記載の発明において、論理構成の観点を示す指標として、文章の接続関係を示す順接及び逆接の接続詞の数、並びに、文章の接続関係を示す順接及び逆接の接続詞の出現パターンのうちの少なくとも一つ以上を用いることで、従来、小論文試験等において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。  
10

**【0090】**

請求項9に記載の発明は、請求項5に記載の発明において、文章の内容について、与えられた出題文との意味的な近さを測定することで評価採点を行なうので、小論文試験において採点に影響を与える様々な誤差要因を排除して、客観的にかつ自動的に評価採点を行なうことができる。

請求項10に記載の発明は、本プログラムをコンピュータに実行させることにより、請求項1～9の何れかの一に記載の発明と同様の作用、効果を奏する。

**【0091】**

請求項11に記載の発明は、記憶しているプログラムにより請求項10に記載の発明と同様の作用、効果を奏する。  
20

**【図面の簡単な説明】**

【図1】本発明の一実施の形態である文章評価採点装置の電気的な接続を示すブロック図である。

【図2】文章評価採点装置を含むネットワークの構成を示す説明図である。

【図3】文章評価採点装置の機能ブロック図である。

【図4】文章評価採点装置が行なう処理のフローチャートである。

【図5】統計量として得た状態遷移の条件付き出力確率の例を示す説明図である。

【図6】文章評価採点装置で示される診断情報の説明図である。  
30

**【符号の説明】**

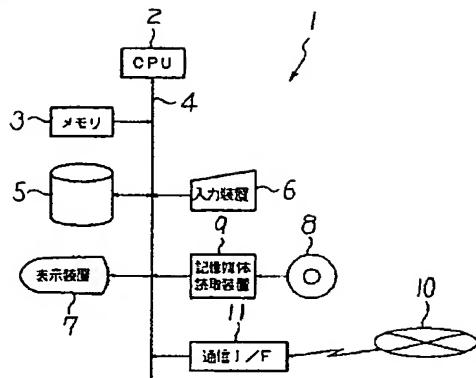
1 文章評価採点装置

S3 設定手段

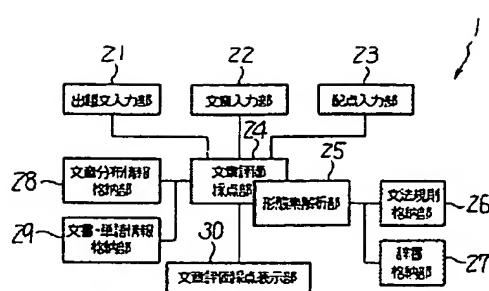
S4～S7 文章評価採点手段

S8 入力手段

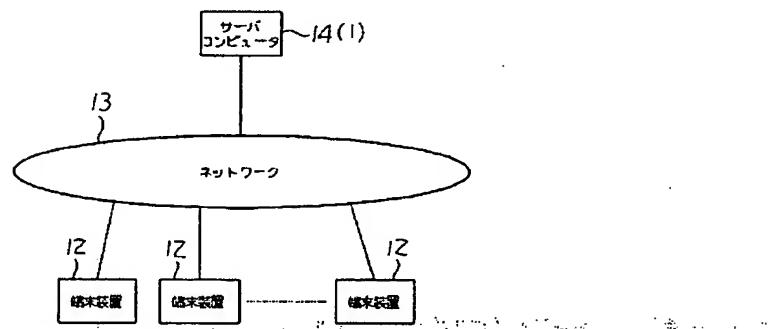
【図 1】



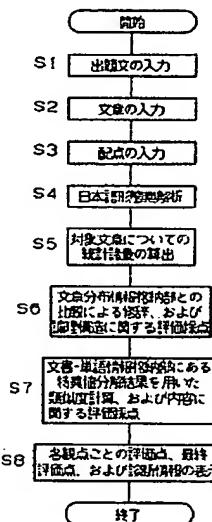
【図 3】



【図 2】



【図 4】



【図 6】

説明情報一覧

「説明に関する評価  
「説明の多様性が不足しています」  
「説明の多様性がやや不足しています」  
「句（読点と句点の間、あるいは読点と句点の間）がはじて（平均的に）長いように見受けられます」  
「句（読点と句点の間、あるいは読点と句点の間）がはじて（平均的に）やや長いように見受けられます」  
「句（読点と句点の間、あるいは読点と句点の間）の長すぎる文があります」  
「句（読点と句点の間、あるいは読点と句点の間）のやや長すぎる文があります」  
「文がはじて（平均的に）長いです」  
「文がはじて（平均的に）少し長いです」  
「句の中の文の数がはじて（平均的に）多いです」  
「句の中の文の数がはじて（平均的に）やや多いです」  
「句の中の文の数が多すぎる大があります」  
「句の中の文の数がやや多すぎる文があります」  
「漢字の使用が少ないように見受けられます」  
「漢字の使用がやや少ないように見受けられます」  
「漢字が非常に使われているように見受けられます」  
「漢字がやや多く以上に使われているように見受けられます」  
「長くて難しい語が少ないように見受けられます」  
「長くて難しい語がやや少ないように見受けられます」  
「受動態の文が全体の分量に比べてやや多いように見受けられます」  
「受動態の文が全体の分量に比べてやや多いように見受けられます」  
「増め込み文が全体の分量に比べてやや多いように見受けられます」  
「増め込み文が全体の分量に比べてやや多いように見受けられます」  
「連用形や受動態の句の並びが多い文が、幾つかあるように見受けられます」  
「連用形や受動態の句の並びの多い文が、ややあるように見受けられます」

【図 5】

(a: 順接, b: 逆接)の状態推移確率

$$\begin{aligned}
 P(a|a) &= 0.48 \quad P(b|a) = 0.52 \quad P(a|b) = 0.36 \quad P(b|b) = 0.64 \\
 P(a|aa) &= 0.35 \quad P(b|ba) = 0.65 \quad P(a|ab) = 0.55 \quad P(b|ab) = 0.44 \\
 P(a|ba) &= 0.28 \quad P(b|ba) = 0.72 \quad P(a|bb) = 0.35 \quad P(b|bb) = 0.65 \\
 P(a|--) &= 0.44 \quad P(b|--) = 0.56
 \end{aligned}$$

内容に関する評価

「説明文との関係が希薄であるように見受けられます」

「説明文との関係がやや希薄であるように見受けられます」

「説明文の説明が不十分であるように見受けられます」

「説明文の説明がやや不十分であるように見受けられます」